**SAMPLE FINAL EXAM.** There are 40 questions. You should be able to answer these questions within three hours; ideally, you can complete the exam within two hours. Good luck.

One simple model for the performance of the S&P500, based on the 607 weeks from 9/17/2001 to 5/6/2013 (see <http://finance.yahoo.com> ), is as follows: each week, there is a 44% chance the market decreases by 1.9%, a 32% chance the market increases by 0.7%, and a 24% chance the market increases by 3.0%. Suppose this describes a probability distribution for the behavior for the S&P500 for foreseeable future. Assume further that the changes from week to week are independent of each other. Assume there are 52 weeks per year.

1. An investor buys an index fund for the S&P500 at $100,000. She uses a “rebalancing” strategy: at the end of each week, she will buy or sell shares so that the balance at the start of the next week is always $100,000. (Assume there are no transaction costs.) Based on the “rebalancing” strategy, what is the chance her net gain after five years will exceed $60,000?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. 0.3% | 1. 1.1% | 1. 2.3% | 1. 16% | 1. 47.5% |

**This is a question about sums of random variables. Each week, given that the gains / losses are applied to a $100,000 balance, the possible net changes are X = { -1900 with probability 0.44, +700 with probability 0.32, and +3000 with probability 0.24 }.**

**E(X) = sum of x\*p = $108 / week**

**SE(X) = sqrt(sum of (x – E(X))2 p) = $1973.205 / week**

**Given five years = 52\*5 = 260 weeks, E(sum) = n E(X) = $28,080 with standard error SE(sum) = sqrt(n)\*SE(X) = $31,817**

**Because n is large and changes are independent from week to week, we can use the normal curve. The chance that the net change exceeds $60,000 is the same as exceeding Z = (60,000 – 28,080) / 31,817 = 1.003, which has a 15.8% chance**

1. An investor buys an index fund for the S&P500 at $100,000. She uses a “rebalancing” strategy: at the end of each week, she will buy or sell shares so that the balance at the start of the next week is always $100,000. (Assume there are no transaction costs.) Based on the “rebalancing” strategy, how many weeks will it be until she has a 90% chance that her net gain will exceed $60,000?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. 20 years | 1. 54 years | 1. 28 years | 1. 11 years | 1. 35 years |

**From above, E(X) = sum of x\*p = $108 / week**

**SE(X) = sqrt(sum of (x – E(X))2 p) = $1973.205 / week**

**Given 28 years = 52\*28 = 1456 weeks, E(sum) = n E(X) = $157,248 with standard error SE(sum) = sqrt(n)\*SE(X) = $75,293. The chance that the net change exceeds $60,000 is the same as exceeding Z = (60,000 – 157,248) / 75,293 = -1.2916, which is the 90% chance desired in the question.**

*Use the following information to answer the next three questions*. A sample of 400 households was asked for their reaction to the Burger King ad that featured the band Coq Roq. Fifty-two (52) percent of the respondents reported a favorable reaction.

1. Set up the appropriate hypotheses that attempt to provide evidence supporting the claim that at least 50% of the households have a favorable reaction.

A) H0: p <0.52 vs HA: p > 0.52 B) H0: p > 0.50 vs HA: p < 0.50 C) H0: p < 0.50 vs HA: p > 0.50

D) H0: p = 0.50 vs HA: p0.50 E) H0: p0.50 vs HA: p = 0.50

**Answer is C**

1. Calculate the appropriate test statistic to test the hypotheses.

A) 0.80 B) 1.60 C) 8.00 D) –1.60 E) –0.80

**Answer is A**

1. Calculate the p-value associated with the test statistic.

A) 0.0001 B) 0.2000 C) 0.2119 D) 0.2881 E) 0.4238

**Answer is C**

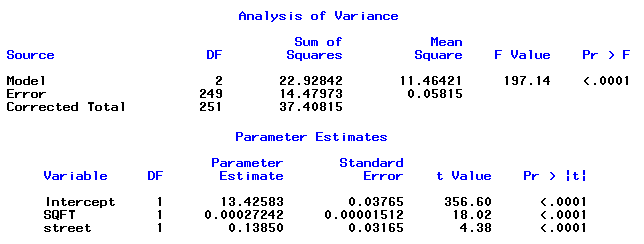
Data are available for 252 houses sold in Manhattan Beach. A researcher records three variables

LOGPRICE natural logarithm of selling price of house

SQFT square footage of house

STREET indicator variable, equal to 1 if address is on a street, 0 otherwise

The researcher performs a multiple regression; the computer output is below.



1. The researcher wonders whether the slope for using SQFT to predict LOGPRICE is the same for houses on streets as it is for houses that are not on streets. Which of the following statements is/are true?
2. In the data, the slope for SQFT is the same for houses on streets as it is for houses not on streets
3. The model assumes the slope for using SQFT to predict LOGPRICE is the same for houses on streets as it is for houses not on streets
4. The F test shows the slope for SQFT is the same for houses on streets as it is for houses not on streets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) (i) only | B) (ii) only | C) (iii) only | D) (ii) and (iii) | E) (i) and (iii) |

**Answer is B, (ii) only**

1. A researcher wants to create a model for employment that assumes that men and women have different initial salaries, but that every year of experience is associated with the same percentage increase in salary. If the response variable is SALARY, and the covariates are YEAR (years of experience) and FEMALE (1 if employee is female, 0 otherwise), what is the appropriate model?
   1. SALARY = b0 + b1 YEAR + b2 FEMALE
   2. ln(SALARY) = b0 + b1 YEAR + b2 YEAR\*FEMALE
   3. ln(SALARY) = b0 + b1 YEAR + b2 FEMALE
   4. SALARY = b0 + b1 YEAR + b2 YEAR\*FEMALE
   5. ln(SALARY) = b0 + b1 YEAR + b2 FEMALE + b3 YEAR\*FEMALE

**Answer is C.**

On March 24, 2013, CBS News asked the equivalent of a simple random sample of 1181 adults nationwide, “Do you think it should be legal or not legal for same-sex couples to marry?” (see <http://www.pollingreport.com/civil.htm> ) The poll broke down the results by age: among n=318 adults who were “18-29”, 72.96% said “legal”; among n=259 adults who were “30-44”, 59.07% said “legal”; among n=270 adults who were “45-64”, 45.93% said “legal”; and among the n=334 who were “65 and older”, 35.03% said legal.

1. Is there statistically significant evidence that the percentage of all adults think it should be legal for same-sex couples to marry is at least 50%? Compute the appropriate (one-sided) p-value.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. p = 1.35% | 1. p = 2.1% | 1. p = 13.4% | 1. p = 14.7% | 1. p = 4.2% |

**Answer is B. This is a test for a single proportion …**

**poverall = (318\*0.7296+259\*0.5907+270\*0.4593+334\*0.3503)/1181 = 0.53**

**Approx SE(poverall) = sqrt[ 0.53\*0.47 / 1181 ]**

1. Is there a statistically significant difference in the percent who say “legal” by age group?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) C = 48.6, compare to 23 df | B) C = 103.4, compare to 23 df | C) C = 52.7, compare to 23 df | D) C = 121.0, compare to 23 df | E) C = 121.0, compare to 21 df |

**Answer is B. This is a chi-square test of independence …**

1. What is a 95% confidence interval for the percentage of adults nationwide who are 18 to 29?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) 68.1%  to 77.8% | B) 22.2%  to 31.9% | C) 47.7%  to 58.6% | D) 23.6%  to 30.3% | E) 24.4%  to 29.5% |

**Answer is E, based on 318 out of 1181 …**

In the Spring 2013 semester of BUAD 310, grades on exam 1 were normally distributed with an average of 67 and a standard deviation of 18, and grades on exam 2 were normally distributed with an average of 59 and a standard deviation of 16. The correlation between exam grades was r = 0.60.

1. Based on the above data, what is the chance that the average of the two exam scores for a randomly chosen student will exceed 70.0?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) 35.9% | B) 32.3% | C) 28.1% | D) 34.0% | E) 18.0% |

**Answer is B. This is a “sum of two correlated random variables” question …**

**Weighted sum = (0.5)(67) + (0.5)(59) = 63**

**SE for sum = 15.2118**

**P(sum > 70.0) = P(Z > [70.0 – 638] / 15.2118) = P(Z > 0.46) = 0.323**

1. Suppose a student scored 59.0 on exam 1. What is the estimated chance that the student scored 67.0 or more on exam 2? Choose the answer closest to correct.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) 30% | B) 10% | C) 20% | D) 40% | E) 50% |

**Answer is A. This is a regression question (using exam 1 to predict exam 2)**

1. Suppose that the final grade in the class is based on the weighted average of three exam scores: score on exam 1 (30% of grade), score on exam 2 (30% of grade), and score on the final exam (40% of grade). Suppose the final has an average of 65 and a standard deviation of 16. Suppose that the correlation between final exam score and score on exam 1 is 0.50, and suppose likewise that the correlation between final exam score and score on exam 2 is also 0.50. What is the chance that a randomly chosen student will have a final grade that exceeds 50.0, the minimum score to get a “C” in BUAD 310?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) 84.2% | B) 86.4% | C) 89.2% | D) 92.4% | E) 96.4% |

**Answer is A. This is a “weighted sum of THREE correlated random variables” question**

**Weighted sum = (0.3)(67) + (0.3)(59) + (0.4)(65) = 63.8**

**SE for sum = 13.7675**

**P(sum > 50.0) = P(Z > [50.0 – 63.8] / 13.7675) = P(Z > -1) = 0.842**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***Use the following to answer the next five questions:***

Data were collected on 140 houses sold in the Rancho Palos Verdes area in California. For each house, four variables were measured:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Description | Correlation  with PRICE | Average | SD |
| PRICE | Selling price of house | 1.000 | 1,682,595 | 652,005 |
| SQFT | Square footage of house | 0.877 | 2687 | 1018 |
| BR | # of bedrooms in house | 0.798 | 3.67 | 0.869 |
| BA | # of bathrooms in house | 0.644 | 2.95 | 1.026 |

The regression equation for predicting PRICE from the three variables was (with standard errors reported in parentheses under the estimated coefficients):

Predicted price = 648170 + 496.645 SQFT + 162,389 BA – 212,074 BR

(152303) (67.375) (60,956) (63,616)

The model fit the data reasonably well, with R2 = 63.1%. The standard error of regression was $400,476.

1. Is the model as a whole statistically significant? Choose the answer closest to correct.

A) No, because F = 1.709

B) No, because standard error of regression is so large

C) Yes, because number of predictors is less than five

D) Yes, because T = 4.26

E) Yes, because F = 77.5

**Answer is E; solve for F by starting with the fact that SSmodel = R2 SStotal and SSresidual = (1 – R2)SStotal, and that k=3 predictors and n=140 (so (n-1)-k = 136)**

1. What is the most reasonable next step to do, given the above model?

A) Drop the Intercept from the model, since it has the largest SE (152303)

B) Drop SQFT from the model, since it has the smallest coefficient

C) Drop BR from the model, since the coefficient is not statistically significant

D) Drop BA from the model, since the coefficient has a negative slope

E) Do nothing, the model is fine as it is

**Answer is E**

1. What is the adjusted R2 for the model? Choose the answer that is closest to correct.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) 58.7% | B) 38.8% | C) 62.3% | D) 78.9% | E) 79.4% |

**Answer is C. Recall adjusted R2 = 1 – (SEY|X / SD(Y))2**

1. Which of the following statements, if any, is/are correct?

i) Converting a bedroom to a bathroom increases selling price by $374,463

ii) Among houses with the same square footage and the same number of

bathrooms, houses with more bedrooms sold for less

iii) The coefficient for bedrooms is not statistically significant

A) Only (i) is correct

B) Only (ii) is correct

C) Only (iii) is correct

D) Both (i) and (ii) are correct

E) All three statements are correct

**Answer is B**

Thirty-one students from BUAD 310 were asked two questions: what score they expected to get on the final (expressed as a percentage), and what the class average would be (also expressed as a percentage). The results are given below:

Estimated score for self: average = 86.2%, SD = 8.9%

Estimated score for class: average = 72.6%, SD = 9.1%

Correlation between “self score” and “class average score”: r = 0.4579

Both “predicted score for self” and “predicted score for class” were close to normally distributed.

1. Based on the data, what percentage of students predicted a higher score for themselves than they did for the class? Choose the answer that is closest to correct.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) 88.5% | B) 92.7% | C) 91.0% | D) 81.2% | E) More than 99.9% |

**Answer is B**

**X = predicted score for self**

**Y = predicted score for class**

**E(X – Y) = 86.2 – 72.6 = 13.6**

**SE(X – Y) = sqrt(8.9^2 – 2(0.4579)(8.9)(9.1) + 9.1^2) = 9.3728**

**P(X – Y > 0) = P(Z > 1.45) = 0.9266**

1. Data were collected for daily return for Bank of America (BAC) for the 2676 trading days from 9/17/2001 to 5/1/2012. The returns had an average of 0.0306%, and 7.78% of the days had returns greater than 3.0%. If daily returns for BAC followed a normal distribution, what percentage of days would have daily returns greater than 4.0%? Choose the answer closest to correct.
2. 1% B) 2% C) 3% D) 4% E) 5%

**Answer is C**

**Average = 0.0306%**

**7.78% had returns > 3% 🡪 3% is 1.42 SD’s above average 🡪 3% is 1.42 SD’s above 0.0306% 🡪 SD = 2.0911%**

**P(return > 4%) = P(Z > (4 – 0.0306)/2.0911) = P(Z > 1.898) = 0.0288 ~ 3%**

1. Data were collected for daily return for Apple (AAPL) for the 2676 trading days from 9/17/2001 to 5/1/2012. A total of 16.6% of the days had returns greater than 2.2%, and 3.0% of the days had returns greater than 5.0%. If returns for AAPL followed a normal distribution, what is the mean for daily return for AAPL? Choose the answer that is closest to correct.
2. 3.07% B) 0.78% C) 1.88% D) –0.97% E) –0.78%

**Answer is E**

**16.6% of days had returns > 2.2% 🡪 2.2% is 0.97 SD’s above average**

**3.0% of days had returns > 5.0% 🡪 5.0% is 1.88 SD’s above average**

**2.2 =  + 0.97 5.0 =  + 1.88  🡪  = 3.0746,  = 0.7826**

Use the following Minitab output for predicting the *natural logarithm* of the sales (*LnSales* in the following) for a sales representative to answer the next eight questions.

The regression equation is

LnSales = 6.76 + 0.0177 YrsExperience + 0.0275 MktPotential + 0.0820 MktShare

Predictor Coef SE Coef T P VIF

Constant 6.7631 0.1474

YrsExperience 0.017680 0.004931 1.400

MktPotential 0.027537 0.005053 1.405

MktShare 0.08196 0.01396 1.036

S = 0.167571 R-Sq = R-Sq(adj) = 79.1%

Analysis of Variance

Source DF SS MS F P

Regression 3.3783 0.000

Residual Error 28

Total

New

Obs Fit SE Fit 95% CI 95% PI

1 8.2636 0.0386 (8.1846, 8.3426) (7.9113, 8.6158)

New

Obs YrsExperience MktPotential MktShare

1 12 20 9

1. What range would you give with 95% confidence for the sales of an individual sales representative with *YrsExperience* = 12, *MktPotential* = 20, and *MktShare* = 9?
   * 1. [7.9113, 8.6158]
     2. [3585.3, 4199.0]
     3. [8.1846, 8.3426]
     4. [2727.9, 5518.2]
     5. [62.5887, 74.2320]

Since we want individual sales representative, we want the 95% prediction interval. Note the 95% prediction interval (PI) for NATURAL LOGARITHM of sales is 7.9113 to 8.6158; therefore, a 95% prediction interval for sales is e7.9113 = 2727.93 to e8.6158 = 5518.16; therefore, the correct answer is (D)

1. What percentage of variation in *LnSales* is explained by the independent variables *YrsExperience*, *MktPotential*, and *MktShare*?
   * 1. 88.9%
     2. 81.1%
     3. 16.8%
     4. 90.1%
     5. 79.1%

“Percentage of variation explained” means unadjusted R2. Intuitively, we know that unadjusted R2 is always larger than adjusted R2, typically by a little bit; this implies the correct answer will be (B).

The formal answer: we know s = sqrt( (1-R2)TSS / (n – k – 1)) = 0.167571 🡪 (1 – R2)TSS = 0.7862. We know SSmodel = R2 TSS = 3.3783, so TSS = 3.3783 + 0.7862 = 4.1645. This implies SD(Y) = sqrt(4.1645 / 31) = 0.3665. We can check our work: based on our calculations, adjusted R2 = 1 – (0.1675712) / (0.36652) = 0.79098, which matches the 79.1% reported in the computer output. Now, R2 = SSmodel / TSS = 3.3783 / 4.1645 = 0.8112, or 81.12%. The correct answer is (B)

1. What are the null and alternative hypotheses for the test of the significance of the independent variable *MktPotential* (*X*2) in the linear regression model?
   * 1. H0: β1 = 0 vs. Ha: β1 ≠ 0
     2. H0: β1 = β2 = β3 = 0 vs. Ha: β1 ≠ 0, β2 ≠ 0, and β3 ≠ 0
     3. H0: β1 = β2 = β3 = 0 vs. Ha: at least one β*j* ≠ 0
     4. H0: β2 = 0 vs. Ha: β2 ≠ 0
     5. H0: β2 ≠ 0 vs. Ha: β2 = 0

Answer is D. The null hypothesis is that, given the other independent variables in the model, the independent variable *MktPotential* is uncorrelated with the Y variable. Since *MktPotential* is the second variable in the multiple regression model, its population coefficient is 2.

1. The value of the total variation is
   * 1. 0.1676
     2. 3.3783
     3. 4.1646
     4. 28
     5. 0.7862

See above. Total variation is TSS = 3.3783 + 0.7862 = 4.1645. The correct answer is (C)

1. The value of the *F*-statistic is
   * 1. 4.2970
     2. 40.10
     3. 1.1261
     4. 5.2971
     5. 0.0250

Here, F = (SSmodel / k) / (SSerror / (n – k – 1)) = (3.3783 / 3) / (0.7862 / 28) = 40.10; answer is (B)

1. The margin of error of the 99% confidence interval (i.e., the “plus or minus” number for the confidence interval) for the slope of the independent variable *MktShare* is
   * 1. 0.028590
     2. 0.013961
     3. 0.038571
     4. 0.407266
     5. 0.013624

For 99% confidence, use T28 = 2.763. Now, MOE = T \* SE = 2.763 \* 0.01396 = 0.03857. Answer is (C)

1. The sample evidence indicates that the population regression equation does not go through the origin at the 0.001 level of significance.
   * 1. True
     2. False
     3. Not sure

The statement is TRUE; this is a question about the intercept, whether 0 = 0. When testing the null hypothesis b0 = 0, the corresponding T-value is (b0 - 0) / SE(b0) = (6.7631 – 0) / 0.1474 = 45.88; the chance of getting a T-value of 45.88 is almost zero (p = 3.39 x 10–28). Answer is thus (A)

1. Suppose you are testing H0: β1 = 0.03 vs. Ha: β1 < 0.03. What should be your decision?
   * 1. Reject H0 at the 0.005 level
     2. Reject H0 at the 0.01 level but not at the 0.005 level
     3. Reject H0 at the 0.05 level but not at the 0.01 level
     4. Reject H0 at the 0.1 level but not at the 0.05 level
     5. Do not reject H0 at the 0.1 level

When testing the null hypothesis b1 = 0.03, the corresponding T-value is (b1 - 1) / SE(b1) = (0.01768 – 0.03) / 0.004931 = -2.498; the chance of getting a T-value of -2.498 is between 0.005 and 0.01 (actual p-value = 0.0093). Answer is thus (B)

1. To better predict the salary using a multiple linear regression model, we would like to include education level, which has the following categories: less than high school, high school graduate, some college, bachelor’s degree, master’s degree, and Ph.D. What will be the effect on the model?
   * 1. Two more independent variables will be included
     2. Three more independent variables will be included
     3. Four more independent variables will be included
     4. Five more independent variables will be included
     5. Six more independent variables will be included

“Education level” is a categorical (ordinal) variable with k = 6 categories, so it can be summarized with k – 1 = 5 categories (e.g., Iless than high school, Ihigh school grad, Isome college, Ibachelor’s, Imaster’s). Answer is (D)

1. Suppose all the points on the scatterplot of Y versus X lie exactly on a line with a negative slope. Also suppose that the standard deviation of the Y values is twice as large as the standard deviation of the X values. Then the slope of the least squares regression line
   1. equals -2
   2. equals 0.5
   3. is negative, but more information is needed to find the exact value
   4. is positive, but more information is needed to find the exact value
   5. None of the above

**Answer is A. Slope of line = r\*SD(Y) / SD(X)**

1. According to a recent survey, about 33% of Americans polled said that they would likely purchase reusable cloth bags for groceries in order to reduce plastic waste. Suppose 45 shoppers are interviewed a local supermarket. What is the probability that more than 12 shoppers will say that they are likely to purchase reusable for groceries?
2. -0.90
3. 0.90
4. 0.267
5. 0.816
6. 0.184

**Answer is D; use binomial distribution**

Use the following information to answer the next questions

Below is a histogram and the five number salary for salaries (in $) for a sample of U.S. marketing managers.



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | Q1 | Median | Q3 | Max |
| 46,360 | 69,693 | 77,020 | 91,750 | 129,420 |

1. The shape of this distribution is
2. symmetric
3. bimodal
4. right skewed
5. left skewed
6. normal

**Answer is C.**

1. The most appropriate measure of central tendency for these data (if one wants to understand the salaries for about half of the managers) is the
2. median
3. mean
4. mode
5. range
6. standard deviation

**Answer is A. The median is the dividing line between the top 50% and the bottom 50%.**

1. The IQR for these data is
2. $83,060
3. $22,057
4. $69,693
5. $77,020
6. $14,566

**Answer is B.**

1. Suppose the marketing manager who was earning $129,420 got a raise and is now earning $140,000. Which of the following statement is true?
2. The mean would increase.
3. The median would increase.
4. The range would increase.
5. Both A and C.
6. All of the above.

**Answer is D.**

1. The table below shows who survived the sinking of the Titanic based on whether they were crew members or passengers booked in first-, second-, or third-class staterooms. To test the hypothesis that the chances of surviving were independent of the status on the ship, we first need to calculate the expected counts under the null hypothesis. What is the expected count for the members of the crew that would not survive the sinking?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Crew* | *First* | *Second* | *Third* |
| Alive | 212 | 202 | 118 | 178 |
| Dead | 673 | 123 | 167 | 528 |

* 1. 599.52
  2. 600
  3. 212
  4. 673
  5. 285.48

**Answer is A.**

1. The internal auditing staff of a local manufacturing company performs a sample audit each quarter to estimate the proportion of accounts that are current. Determine the sample size needed in order to be 95% confident that the sample proportion of the current customer accounts is within 0.015 of the true proportion of all current accounts for this company.
   1. 4,268
   2. 4,269
   3. 3,007
   4. 3,006
   5. 1,068

**Answer is B (you need to round up to get the 95% confidence level)**

1. Suppose you obtain a simple random sample of 50 working LA residents, and you want to test whether the distribution of job satisfaction level (high, medium, and low) stays the same across age groups (younger than 25, 25-45, and 45+). Which distribution will you use to find the p-value of the test?
   1. standard normal
   2. t-distribution with 2 df
   3. t-distribution with 49 df
   4. chi-square distribution with 2 df
   5. chi-square distribution with 4 df

**Answer is E**

1. Which of the following is true?
   1. Confidence intervals are wider for 88% confidence level than for 84% confidence level.
   2. All chi-square distributions have the mean of zero.
   3. All t distributions are skewed to the right.
   4. All of the above
   5. None of the above

**Answer is A**

1. Life insurance. According to the current Commissioners’ Standard Ordinary mortality table, adopted by stats insurance regulators in December 2002, a 25-year-old man has these probabilities of dying during the next five years:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Age at death | 25 | 26 | 27 | 28 | 29 |
| Probability | 0.00039 | 0.00044 | 0.00051 | 0.00057 | 0.0006 |

An online insurance site offers a term insurance policy that will pay $100,000 if a 25-year-old man dies within the next 5 years. The cost is $175 per year, due on the individual’s birthday. So the insurance company will take in $875 from this policy if the man does not die within five years. If he does die, the company must pay $100,000; so if the individual dies the day after his 25th birthday, the company will have a net loss of $99,825 ($175 payment less $100,000 payment).

What is the insurance company’s mean cash intake (= INCOME – LOSS) and standard deviation from such policies?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A) mean = $175.49, SD = $8343.48 | B) mean = $448.22, SD = $5019 | C) mean = $0.49, SD = $8343 | D) mean = $623.22, SD = $5019 | E) mean = -$75.03, SD = $4984 |

**Answer is D. The loss depends on how many premiums were paid by the individual. If the individual survives, the insurance company earns $875; otherwise, the payouts are as follows:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Age at death | 25 | 26 | 27 | 28 | 29 |
| Loss | -$99,825 | -$99,650 | -$99,475 | -$99,300 | -$99,125 |

**Now compute expected value (the mean) as “sum of (outcome\*probability) = 0.00039\*(-99825) + 0.00044\*(-99650) + … + (0.00060)\*(-99125) + 0.99\*(875)” = $623.22. Compute SE as “square root of [sum of (outcome – mean)2(probability) ] “ = 5019.308**

**END SAMPLE FINAL EXAM**